Math-Net.Ru

Общероссийский математический портал

В. П. Боровиков, Об асимптотическом распределении статистик, основанных на выборочных промежутках, с оцененными параметрами, *Теория вероятн. и ее примен.*, 1996, том 41, выпуск 1, 181–186

DOI: <https://doi.org/10.4213/tvp2791>

Использование Общероссийского математического портала Math-Net.Ru подразумевает, что вы прочитали и согласны с пользовательским соглашением
<http://www.mathnet.ru/rus/agreement>

Параметры загрузки:

IP: 3.85.5.30

26 апреля 2023 г., 14:52:33

В силу правила Лопиталя

$$\lim_{p \rightarrow 1/2} \frac{(1-p)^t - p^t}{1-2p} = t 2^{1-t}.$$

Поэтому из (13) легко получаем

$$\lim_{p \rightarrow 1/2-0} \Delta(p) = \frac{2^{t-1}}{t+1}.$$

Очевидно, отсюда следует (7). Теорема 2 доказана.

СПИСОК ЛИТЕРАТУРЫ

1. Эбралидзе Ш. С. Неравенства для вероятностей больших отклонений в терминах псевдомоментов. — Теория вероятн. и ее примен., 1971, т. XVI, в. 4, с. 760–765.
2. Боровков А. А. Замечания о неравенствах для сумм независимых величин. — Теория вероятн. и ее примен., 1972, т. XVII, в. 3, с. 588–590.
3. Боровков А. А. О скорости сходимости в принципе инвариантности. — Теория вероятн. и ее примен., 1973, т. XVIII, в. 2, с. 217–234.
4. Петров В. В. Предельные теоремы для сумм независимых случайных величин. М.: Наука, 1987, 317 с.
5. Проторов Ю. В. Сходимость случайных процессов и предельные теоремы теории вероятностей. — Теория вероятн. и ее примен., 1956, т. I, в. 2, с. 177–238.

Поступила в редакцию
2.VIII.1995

© 1996 г.

БОРОВИКОВ В. П.*

ОБ АСИМПТОТИЧЕСКОМ РАСПРЕДЕЛЕНИИ СТАТИСТИК, ОСНОВАННЫХ НА ВЫБОРОЧНЫХ ПРОМЕЖУТКАХ, С ОЦЕНЕННЫМИ ПАРАМЕТРАМИ

Установлено, что предельное (при $n \rightarrow \infty$) распределение статистик вида

$$G_n(\theta) = n^{-1/2} \sum_{1 \leq i \leq n} g(n f(X_{(i)}, \hat{\theta}_n) \Delta X_i),$$

где $X_{(i)}$ есть i -я порядковая статистика, ΔX_i есть i -й выборочный промежуток, $\hat{\theta}$ — оценка неизвестного параметра θ , не зависит от θ при нулевой гипотезе.

Описан широкий класс альтернатив, для различения которых естественно применять статистики, основанные на выборочных промежутках.

Ключевые слова и фразы: выборочный промежуток или спейсинг, порядковые статистики, предельное распределение, оценка параметра.

*Институт радиотехники и электроники РАН, ул. Моховая, 11, 103907 Москва, Россия.

1. Изучение статистик, основанных на спейсингах, имеет давнюю историю, см., в частности, работы Фишера, Дарлингга, Леви, Ле Кама, Чибисова, Пайка, Кудлаева [1]–[7].

Основной факт, устанавливаемый нами, состоит в том, что предельное распределение статистик, основанных на выборочных промежутках и с оцененными параметрами, не зависит от истинного значения этих параметров и, вообще, от оценивания параметров, т.е. совпадает с распределением, когда параметры известны. В работе [8] установлен факт независимости предельного распределения от оцениваемых параметров для одной специальной статистики, основанной на выборочных промежутках.

Для других классических статистик, например, статистики Колмогорова дело обстоит не так просто. В частности, потребовались значительные усилия, завершившиеся результатами Хмаладзе [9], чтобы построить критерий, основанный на статистике Колмогорова, обладающий свойством свободной распределенности.

2. Перейдем к точным формулировкам и доказательству. Пусть $X_1, \dots, X_n$ — независимые одинаково распределенные случайные величины (с.в.), каждая из которых имеет функцию распределения (ф.р.) F с плотностью f . Пусть относительно плотности f проверяется гипотеза $H_0: f \in \{f(\cdot, \theta)\}$, где $\theta \in \Theta$, Θ есть некоторый компакт в $\mathbf{R}_k$, $k \geq 1$.

Пусть $X_{(1)} \leq X_{(2)} \leq \dots \leq X_{(n)}$ есть вариационный ряд, построенный по с.в. $X_1, \dots, X_n$. Расстояния между соседними членами вариационного ряда есть выборочные промежутки или, более кратко, спейсинги $\Delta X_i = X_{(i+1)} - X_{(i)}$, $1 \leq i < n$.

Определим также с.в. $\Delta F(X_{(i)}, \theta) = F(X_{(i+1)}, \theta) - F(X_{(i)}, \theta)$, где $F(\cdot, \theta)$ есть ф.р. с.в. X_i при гипотезе H_0 .

Рассмотрим следующие функции, основанные на спейсингах:

$$G_n(\theta) = n^{-1/2} \sum_{1 \leq i < n} g\left(n f(X_{(i)}, \theta) \Delta X_i\right),$$

$$G'_n(\theta) = n^{-1/2} \sum_{1 \leq i < n} g\left(n \Delta F(X_{(i)}, \theta)\right),$$

где g — некоторая измеримая функция, $g: \mathbf{R}_+ \rightarrow \mathbf{R}_1$.

Далее мы будем рассматривать асимптотическое, при $n \rightarrow \infty$, поведение статистик $G_n(\hat{\theta}_n)$ и $G'_n(\hat{\theta}_n)$ при гипотезе H_0 , когда вместо истинного значения параметра θ , подставлена его состоятельная оценка $\hat{\theta}_n$.

Отметим, что нормировка в определении $G_n(\theta)$ и $G'_n(\theta)$ выбрана так, чтобы при естественном центрировании эти статистики имели невырожденное предельное распределение при гипотезе H_0 . (В качестве центрирующих констант при этом следует взять $Eg(n\Delta U_i)$, где ΔU_i есть i -й спейсинг, построенный по равномерной выборке, $1 \leq i < n$.)

Сформулируем ряд условий относительно функций f и g .

Для краткости далее рассматриваем случай одномерного параметра θ . Все результаты естественным образом распространяются и на случай многомерного параметра.

(f1) Функция f дифференцируема по x и дважды дифференцируема по θ , $f'_x, f''_{\theta\theta}$ равномерно ограничены.

$$(f2) \quad \frac{\partial}{\partial \theta} \int f(x, \theta) dx = \int \frac{\partial f}{\partial \theta}(x, \theta) dx = 0.$$

$$(f3) \quad \text{supp } f(\cdot, \theta) = [0, 1], f(x, \theta) > \delta > 0, x \in [0, 1], \theta \in \Theta.$$

(g1) Функция g дважды непрерывно дифференцируема и при больших x имеет место неравенство $|g''(x)| \leq \exp\{\varepsilon \delta x\}$, где $\varepsilon \in (0, \frac{1}{2})$, а δ взято из условия (f3).

Сформулируем условие регулярности.

(А) Имеют место условия (f1), (f2), (f3), (g1), оценка $\hat{\theta}_n$ такова, что $|\hat{\theta}_n - \theta| = o_P(n^{-1/4})$, $n \rightarrow \infty$.

Теорема. Пусть выполнено условие (А). Тогда при гипотезе H_0 справедливы соотношения:

$$1) G_n(\hat{\theta}_n) - G_n(\theta) = o_P(1),$$

$$2) G'_n(\hat{\theta}_n) - G'_n(\theta) = o_P(1).$$

Доказательство. При доказательстве используется метод работы [10]. Ограничимся лишь доказательством части 1). Прежде всего, в силу формулы Тейлора, имеем

$$G_n(\hat{\theta}_n) = G_n(\theta) + (\hat{\theta}_n - \theta) n^{-1/2} \sum_i g'_{\theta_i} f'_{\theta_i} n \Delta X_i + \frac{(\hat{\theta}_n - \theta)^2}{2!} \times n^{-1/2} \sum_i \left[\bar{g}''_{\theta_i} (\bar{f}'_{\theta_i})^2 (n \Delta X_i)^2 + \bar{g}'_{\theta_i} \bar{f}''_{\theta_i} n \Delta X_i \right] = G_n(\theta) + \varepsilon_{1n} + \varepsilon_{2n},$$

где $g_{\theta_i} = g(f(X_{(i)}, \theta) n \Delta X_i)$, $f_{\theta_i} = f(X_{(i)}, \theta)$, $\bar{g}_{\theta_i}$, $\bar{g}''_{\theta_i}$, $\bar{f}'_{\theta_i}$, $\bar{f}''_{\theta_i}$ есть значения соответствующих функций в промежуточной точке $\bar{\theta}_n \in (\hat{\theta}_n \wedge \theta, \hat{\theta}_n \vee \theta)$.

Рассмотрим с.в. ε_{1n} . Введем последовательность σ -алгебр $\mathcal{F}_i^n = \sigma(X_{(1)}, \dots, X_{(n)})$, $1 \leq i < n$, $n \geq 1$. Воспользуемся разложением второго сомножителя в ε_{1n} на «компенсатор» и «мартингальную составляющую» (см. [10], [11]): $\varepsilon_{1n} = (\hat{\theta}_n - \theta)(m_n + k_n)$, где

$$m_n = n^{-1/2} \sum_i \left(g'_{\theta_i} f'_{\theta_i} n \Delta X_i - f'_{\theta_i} E(g'_{\theta_i} n \Delta X_i | \mathcal{F}_i^n) \right),$$

$$k_n = n^{-1/2} \sum_i f'_{\theta_i} E(g'_{\theta_i} n \Delta X_i | \mathcal{F}_i^n).$$

Стандартные предельные теоремы для мартингалов (см., например, [12]) легко дают, что $m_n = o_P(1)$, $n \rightarrow \infty$.

Далее, используя тот факт, что (см. [10], [11])

$$E(g'_{\theta_i} n \Delta X_i | \mathcal{F}_i^n) = f_{\theta_i}^{-1} \int_0^\infty g'(x) x e^{-x} dx + o_P(n^{-1/2}),$$

можно показать, что с.в. k_n аппроксимируются величинами l_n вида

$$l_n = n^{-1/2} \sum_i \frac{f'(X_{(i)}, \theta)}{f(X_{(i)}, \theta)} \int_0^\infty g'(x) x e^{-x} dx = o_P(1), \quad n \rightarrow \infty.$$

Именно, $k_n - l_n = o_P(1)$, $n \rightarrow \infty$, следовательно, в силу состоятельности оценки $\hat{\theta}_n$, $\varepsilon_{1n} = o_P(1)$, $n \rightarrow \infty$.

Далее, в силу условия $n^{1/4}(\hat{\theta}_n - \theta) = o_P(1)$ и условий (f1)–(f3), (g1),

$$\varepsilon_{2n} = o_P(1), \quad n \rightarrow \infty.$$

Действительно,

$$\varepsilon_{2n} = \frac{1}{2!} n^{1/2} (\bar{\theta}_n - \theta)^2 \bar{\varepsilon}_{2n}, \quad \bar{\varepsilon}_{2n} = o_P(1), \quad n \rightarrow \infty,$$

так как условная ф.р. i -го выборочного промежутка относительно σ -алгебры $\mathcal{F}_i^n$, имеющая вид

$$P\{n \Delta X_i > x | \mathcal{F}_i^n\} = \left(1 - \frac{F(X_{(i)} + x/n) - F(X_{(i)})}{1 - F(X_{(i)})} \right)^{n-i},$$

аппроксимируется экспоненциальной ф.р. со случайным параметром (см. [10], [11]) и при больших x для функции g'' и, очевидно, для g' имеет место неравенство условия (g1).

Собирая вместе полученные выкладки, приходим к утверждению теоремы в части 1).

Аналогично доказывается часть 2).

3. Некоторые замечания и дополнительные рассуждения.

1. В условиях теоремы имеет место близость статистик $G_n(\hat{\theta}_n)$ и $G'_n(\hat{\theta}_n)$. Более точно, $G_n(\hat{\theta}_n) - G'_n(\hat{\theta}_n) = o_P(1)$, $n \rightarrow \infty$.

2. При выполнении специальных условий на функцию g теорема имеет место и при весьма медленной сходимости оценок. Сформулируем эти условия.

(g2) Для целого $k > 1$ функция g является $k + 1$ раз непрерывно дифференцируемой, причем имеют место равенства

$$\int g''(x) x^2 e^{-x} dx = \dots = \int g^{(k)}(x) x^k e^{-x} dx = 0,$$

и при больших x и некотором $\epsilon \in (0, \frac{1}{2})$ имеет место неравенство

$$|g^{(k+1)}(x)| \leq \exp\{\epsilon \delta x\}.$$

Пусть дополнительно к (f1) плотность f также $k + 1$ раз дифференцируема по θ и $f_{\theta}^{(k+1)}$ равномерно ограничена, а оценка $\hat{\theta}_n$ такова, что

$$(\hat{\theta}_n - \theta)^{k+1} = o_P(n^{-1/2}), \quad k > 1, \quad n \rightarrow \infty.$$

3. Для частичных сумм функции g от спейсингов вида

$$X_n(t) = n^{-1/2} \sum_{i < nt} g(nf(X_{(i)}, \hat{\theta}_n) \Delta X_i),$$

а также

$$X'_n(t) = n^{-1/2} \sum_{i < nt} g(n\Delta F(X_{(i)}, \hat{\theta}_n)), \quad t \in [0, 1],$$

имеет место аналогичный результат, если в дополнение к условию (g1) потребовать выполнения условия

$$(g3) \quad \int_0^\infty g'(x) x e^{-x} dx = 0.$$

4. Если оценка $\hat{\theta}_n$ является $n^{1/2}$ -состоятельной, то для справедливости теоремы для частичных сумм относительно функции g достаточно требовать лишь выполнения условия (g3), которое можно назвать условием асимптотической параметрической несмещенности.

5. Статистики, основанные на спейсингах, естественно применять для проверки гипотез, высказанных относительно плотностей (а не функций распределения). Можно указать достаточно широкий класс альтернатив, которые различимы с помощью критериев, основанных на спейсингах, но не различимы классическими критериями, основанными на эмпирической функции распределения. В частности, такими альтернативами являются следующие.

Рассмотрим случай простой гипотезы $H_0: f(t) = f_0(t) = 1, t \in [0, 1]$. В качестве альтернативных к H_0 рассмотрим гипотезы

$$H_n: f_n(t) = 1 + a_n \cos(2\pi M_n t),$$

где M_n есть некоторая последовательность чисел, $a_n = o(1)$, $n \rightarrow \infty$.

Для тестирования H_0 против H_n воспользуемся статистиками G'_n , полагая, например, $g(x) = x^2$. Рассмотрим с.в.

$$g'_n = n^{-1/2} \sum_{i \leq n} [(n \Delta X_i)^2 - E_0(n \Delta X_i)^2],$$

которые при гипотезе H_0 сходятся по распределению к нормальной с.в. со средним 0 и дисперсией, вычисляемой по методу [6] или [10].

При альтернативах H_n с.в. g'_n имеют также нормальное распределение, однако в математическом ожидании возникает сдвиг γ_n вида

$$\gamma_n = n^{-1/2} \sum_i [E_n(n \Delta X_i)^2 - E(n \Delta U_i)^2],$$

где E_n обозначает усреднение при альтернативах.

Вычислим данный сдвиг, заметив, прежде всего, что

$$E_n(n \Delta X_i)^2 = \frac{2}{(f_n(F_n^{-1}(i/n)))^2} + o(1), \quad n \rightarrow \infty,$$

следовательно,

$$\gamma_n \sim 2n^{1/2} \left(\int_0^1 \frac{dt}{(f_n(F_n^{-1}(t)))^2} - 1 \right) = 2n^{1/2} \left(\int_0^1 \frac{d\tau}{f_n(\tau)} - 1 \right).$$

Вспоминая выражение для f_n , легко видеть, что $\gamma_n \sim n^{1/2} a_n^2$, $n \rightarrow \infty$.

Таким образом, альтернативы H_n отличимы от гипотезы с помощью статистик g'_n при условии $\gamma_0 = \lim_{n \rightarrow \infty} n^{1/2} a_n^2 \geq 0$. Пусть теперь выполнено условие $M_n^{-1} = o(n^{-1/2})$. Легко видеть, что при данном условии ни критерий Колмогорова, ни ω^2 не различают H_n .

Естественное обобщение данных альтернатив приводит к классу плотностей, допускающих представление вида

$$f_n(t) = 1 + \sum_{1 \leq l \leq k} a_{ln} \cos(2\pi M_{ln} t),$$

с условием

$$\lim_{n \rightarrow \infty} n^{1/2} (a_{1n}^2 + \dots + a_{kn}^2) > 0, \quad \exists M_{l'n}: M_{l'n}^{-1} = o(n^{-1/2}).$$

Автор глубоко благодарен проф. Д. М. Чибисову за критические замечания, улучшившие статью.

СПИСОК ЛИТЕРАТУРЫ

1. Fisher R. Contribution to Mathematical Statistics. New York: Wiley, 1950.
2. Darling D. On a class of problems related to the random division of an interval. — Ann. Math. Statist., 1953, v. 24, № 1, p. 239–253.
3. Lévy P. Sur la division d'un segment par des points choisis au hasard. — C. R. Acad. Sci. Paris, 1939, v. 208, p. 147–149.
4. Le Came L. Une theoreme sur la division d'un intervalle par des points pris au hasard. — Publ. Inst. Stat. Univ. Paris, 1958, v. 7, p. 7–16.

5. Чибисов Д. М. О критериях согласия, основанных на выборочных промежуточных. — Теория вероятн. и ее примен., 1961, т. 6, в. 3, с. 354–358.
6. Pyke R. Spacings revisited. — In: Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability. Vol. 1. Berkeley-Los Angeles: Univ. California Press, 1971, p. 417–427.
7. Кудлаев Э. М. Асимптотическая нормальность статистик некоторых классов критериев согласия и однородности. — Докл. АН СССР, 1978, т. 240, № 3, с. 534–537.
8. Cheng R. C. H., Stephens M. A. A goodness-of-fit test using Moran's statistics with estimated parameters. — Biometrika, 1989, v. 76, № 2, p. 385–392.
9. Хмаладзе Э. В. Некоторые применения теории мартингалов в статистике. — Успехи матем. наук, 1982, т. 37, в. 6, с. 193–212.
10. Боровиков В. П. Некоторые предельные теоремы для статистик, являющихся частичными суммами функций от спейсингов. — Теория вероятн. и ее примен., 1987, т. 32, в. 1, с. 92–104.
11. Borovikov V. P. On the theory of general spacings, Part 1. — Mitteilungen aus dem Mathem. Seminar Giessen, Sonderdruck aus Heft 215, Giessen 1994, p. 43–77.
12. Ширяев А. Н. Вероятность. М.: Наука, 1980, 575 с.

Поступила в редакцию
18.XI.1993

© 1996 г.

ДЪЯКОНОВА Е. Е.*

БЛИЗКИЕ К КРИТИЧЕСКИМ ВЕТВЯЩИЕСЯ ПРОЦЕССЫ С МИГРАЦИЕЙ¹⁾

В предлагаемой работе изучается ветвящийся процесс Гальтона–Ватсона с однородными по времени и не зависящими от состояния процесса иммиграцией и эмиграцией. Показано, что при некоторых условиях, наложенных на распределение числа мигрирующих частиц и распределение числа потомков, специальным образом нормированное стационарное распределение этого процесса сходится при $t \nearrow 1$ к равномерному распределению на отрезке $[0, 1]$. Здесь t — среднее значение числа потомков одной частицы в следующем поколении.

Ключевые слова и фразы: ветвящийся процесс Гальтона–Ватсона, докритический процесс, переходные явления, предельное распределение.

1. Введение. Описание модели. Рассмотрим следующий процесс Гальтона–Ватсона с миграцией. Предположим, что заданы два независимых семейства целочисленных случайных величин $\xi_k^{(n)}$, η_n , $k = 0, 1, 2, \dots$, $n = 0, 1, \dots$. Пусть случайные величины η_n , $n = 0, 1, \dots$, одинаково распределены и

$$P\{\eta_n = i\} = p_i, \quad i = 0, 1, \dots, \quad P\{\eta_n = -j\} = q_j, \quad j = 1, \dots, T, \\ \sum_{i=0}^{\infty} p_i + \sum_{j=1}^T q_j = 1, \quad (1)$$

*Математический институт им. В. А. Стеклова РАН, ул. Вавилова, 42, Москва ГСП-1 117966, Россия.

¹⁾Работа выполнена при финансовой поддержке Российского фонда фундаментальных исследований, проект № 93-0111443.